



# The development of a regional geomagnetic daily variation model using neural networks

P. R. Sutcliffe

## ► To cite this version:

P. R. Sutcliffe. The development of a regional geomagnetic daily variation model using neural networks. Annales Geophysicae, 2000, 18 (1), pp.120-128. hal-00316553

**HAL Id: hal-00316553**

**<https://hal.science/hal-00316553>**

Submitted on 1 Jan 2000

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The development of a regional geomagnetic daily variation model using neural networks

P. R. Sutcliffe

Hermanus Magnetic Observatory, PO Box 32, Hermanus 7200, South Africa

Received: 5 January 1999 / Revised: 28 June 1999 / Accepted: 20 July 1999

**Abstract.** Global and regional geomagnetic field models give the components of the geomagnetic field as functions of position and epoch; most utilise a polynomial or Fourier series to map the input variables to the geomagnetic field values. The only temporal variation generally catered for in these models is the long term secular variation. However, there is an increasing need amongst certain users for models able to provide shorter term temporal variations, such as the geomagnetic daily variation. In this study, for the first time, artificial neural networks (ANNs) are utilised to develop a geomagnetic daily variation model. The model developed is for the southern African region; however, the method used could be applied to any other region or even globally. Besides local time and latitude, input variables considered in the daily variation model are season, sunspot number, and degree of geomagnetic activity. The ANN modelling of the geomagnetic daily variation is found to give results very similar to those obtained by the synthesis of harmonic coefficients which have been computed by the more traditional harmonic analysis of the daily variation.

**Key words.** Geomagnetism and paleomagnetism (time variations; diurnal to secular) · Ionosphere (modelling and forecasting)

---

## 1 Introduction

Global and regional geomagnetic field models are routinely utilised in a variety of fields. For example, they find practical application in magnetic navigation, orientation control, and geophysical exploration and are also widely used by researchers in geophysics and space physics. In a number of fields of endeavour, users are tending to increasingly employ geomagnetic field models and to require models of greater accuracy and resolution (Kerridge, 1993; Newitt, 1993). Certain users have

recently expressed interest in the availability of geomagnetic field models which, in addition to the main field, also make allowance for short term geomagnetic field variations.

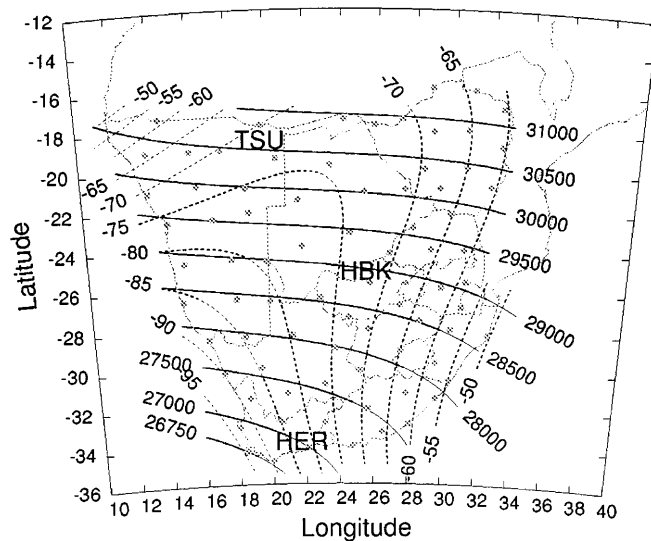
Over the past few years artificial intelligence (AI) methods have been increasingly recognised as powerful analysis tools in solar-terrestrial physics (McPherron, 1993; Joselyn *et al.*, 1993). Artificial neural networks (ANNs) are a branch of AI methods which are proving particularly successful in solar-terrestrial time series prediction and pattern recognition; they appear to be especially effective in modelling the time development of irregular processes (Koons and Gorney, 1991; Lundstedt, 1992; Gorney *et al.*, 1993; Lundstedt and Wintoft, 1994; Willisroft and Poole, 1996; Wu and Lundstedt, 1996; Sutcliffe, 1997; Weigel *et al.*, 1999).

A limitation of most geomagnetic main field models is that they do not take account of short term variations in the geomagnetic field, such as the regular daily variation. Separate models have been developed to describe the solar quiet day (Sq) variation (Campbell and Schiffmacher, 1985, 1988); however, the harmonic (Fourier) analysis (HA) method traditionally used for this purpose is unable to model the irregular part of this variation. Here we describe for the first time the development of a geomagnetic daily variation model utilising artificial neural networks. The objective was to investigate whether ANN techniques can help to overcome the limitations mentioned. The model developed was for the southern African region; however, the method used could be applied to any other region or even globally.

## 2 Geomagnetic field models

The purpose of a geomagnetic field model is to provide the user with estimates of the values of a selected component of the Earth's magnetic field as functions of position and epoch. The most commonly used global field model is the International Geomagnetic Reference Field (IGRF) (Barton, 1997). The Hermanus Magnetic

Observatory (HMO), like similar institutions in other countries, generates a number of regional geomagnetic field models. The map of southern Africa in Fig. 1 shows the positions of the three continuous recording stations HER, HBK, and TSU (see Table 1 for a listing of their coordinates) as well as the field stations used for modelling purposes. Also shown are the iso-contours of the total field  $F$  for epoch 1995 for a third degree polynomial model for the southern African region (Kotzé, 1992). The solid contours give the geomagnetic total field intensities as functions of position, while the dashed contours indicate the annual change of field values. However, neither of these models provide the user with estimates of short term variations in the geomagnetic field, such as regular daily variations or geomagnetic disturbances. A number of more comprehensive models, which enable corrections to be made for certain magnetospheric and ionospheric field variations have been described in the literature (Langel *et al.*, 1996; Purucker *et al.*, 1997). The model described is intended as an extension to commonly used field models for users who wish to take the regular geomagnetic daily variation into account.



**Fig. 1.** Map of southern Africa showing the positions of the three continuous recording stations, *HER*, *HBK*, and *TSU*, and the field stations used for deriving the HMO's geomagnetic field models. Also shown are the iso-contours of the geomagnetic field model for the total field  $F$  for epoch 1995; *solid contours* give the field intensities, while the *dashed contours* indicate the secular variation

**Table 1.** Names, codes, and position coordinates of the observing stations whose data were used in the geomagnetic daily variation model

Name	Code	Geographic co-ordinates		Geomagnetic latitude
		Latitude	Longitude	
Hermanus	HER	34.425°S	19.225°E	-42.49
Hartbeesthoek	HBK	25.882°S	27.701°E	-36.46
Tsumeb	TSU	19.200°S	17.583°E	-30.98

### 3 Geomagnetic field variations

Inspection of a continuous recording of any of the components of the geomagnetic field typically reveals two types of variations (Parkinson, 1983). Firstly, recordings at non-polar latitudes will almost always exhibit a smooth regular variation, known as the solar quiet day or Sq variation, which arises as the magnetic signature of E-region ionospheric currents driven by dynamo action (Campbell, 1989). Secondly, recordings sometimes exhibit rapid irregular fluctuations referred to as geomagnetic disturbances or storms, the magnitude of which may be such that the regular Sq variation is swamped and thus not easily discernable. Although the Sq variation is the most regular of all the geomagnetic field variations, tending to repeat itself with a periodicity of 24 h, significant day-to-day differences do occur (Hibberd, 1981).

Consideration of spectra representing the Sq variation indicates that the power of the Sq variation is strongly concentrated in a number of narrow frequency bands at 1, 2, 3, and sometimes more cycles per day. This suggests that a good representation of the Sq variation for a particular location and date should be yielded by the synthesis of the first 3 or 4 harmonic components of a 24 h wave. A representation of the total field  $F$ , for example, can thus be expressed mathematically as:

$$F = F_o + \Delta F = F_o + \sum_{n=1}^N F_n \sin(nt + f_n) + \epsilon \quad (1)$$

where  $F_o$  is the mean magnetic field value over a day,  $\Delta F$  is the total Sq variation,  $F_n$  and  $f_n$  are the amplitude and phase respectively of the  $n$ th harmonic component of  $\Delta F$ , and  $\epsilon$  is the irregular part of  $\Delta F$ . Similar expressions will hold for the other geomagnetic field components. Note that the traditional geomagnetic field models, such as IGRF, provide estimates of  $F_o$  as functions of epoch and position. A daily variation model, on the other hand, would be expected to provide an estimate of  $\Delta F$  for the prevailing conditions.

### 4 Reasons for a neural network based geomagnetic daily variation model

Harmonic analysis (HA)-based methods have traditionally been used to derive coefficients to represent the Sq variation (Malin and Chapman, 1970); however, they suffer a major shortcoming and can be computationally involved. The ANN-based techniques which we have used for the development of a geomagnetic daily variation model should contribute to overcoming these shortcomings in the following ways:

1. Sq variations synthesised from HA determined harmonic coefficients, that is from the  $F_n$  and  $f_n$  in Eq. (1), provide reasonable representations of the observed Sq variations for individual days. However, a major shortcoming of HA-based methods is that they are not able to account for the observed day-to-day

variability of the Sq variation; that is, HA-based methods are not able to provide an estimate of the irregular part of the Sq variation as represented by  $\epsilon$  in Eq. (1). ANN-based techniques, on the other hand, have been particularly successful in predicting a variety of irregular magnetospheric processes such as magnetic storms and disturbance indices. The motivation for the inclusion of the degree of geomagnetic activity as a parameter in our ANN-based geomagnetic daily variation model was that it may partially account for the day-to-day variability of the daily variation.

2. Previous studies have shown that the amplitude  $F_n$  and phase  $f_n$  exhibit strong dependencies on season and solar cycle and a weaker dependence on the degree of geomagnetic activity (Malin *et al.*, 1975; Sutcliffe, 1977). Consequently, when the harmonic coefficients which represent the Sq variation at a particular location are derived, it is necessary to consider these dependencies and to group the data for specific sets of geophysical conditions. As a result, a separate set of harmonic coefficients must be determined for each of many sets of conditions. Furthermore, the harmonic coefficients are latitude- and, to a lesser extent, longitude-dependent. For example, Campbell and Schiffmacher (1985, 1988) used harmonic analysis in the development of an analytical representation of the global and seasonal Sq variation fields for sunspot minimum and magnetically quiet conditions. Their method thus required that a separate set of harmonic coefficients be determined for each month of a quiet year at each of a global distribution of magnetic observatories. These coefficients were then combined using a somewhat complex process of synthesis to provide an Sq model with seasonal and latitudinal dependence for each of a number of longitudinal regions in the Northern and Southern Hemispheres. An ANN-based model is simpler in that it uses only one set of data, which are functions of all the input parameters, resulting in one set of coefficients (weights).

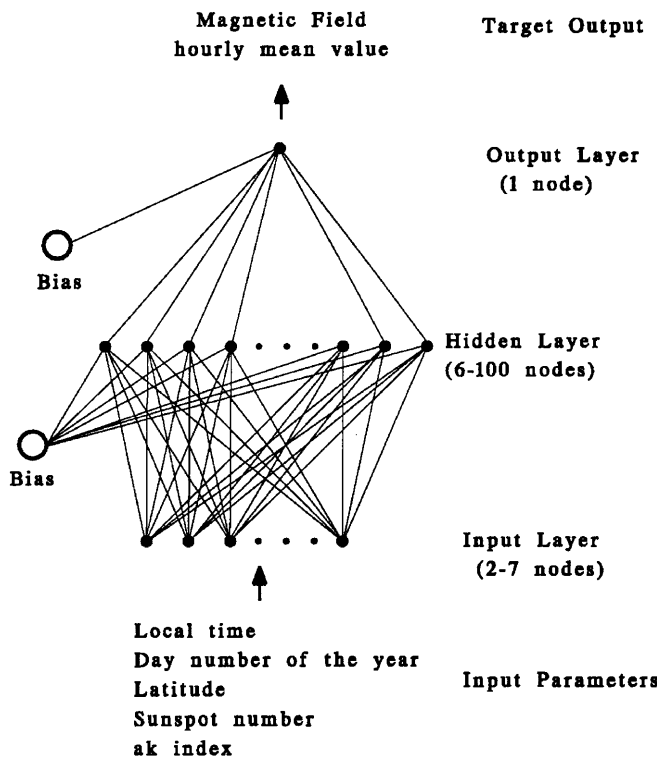
3. As a result of their data selection procedure, the Campbell and Schiffmacher (1985, 1988) model is valid only for very quiet conditions and consequently does not take account of sunspot number or geomagnetic activity. Harmonic coefficients representing the sunspot cycle could, of course, be determined by repeating their analyses at different stages of the sunspot cycle. Olsen (1993), e.g., investigated the solar cycle dependence of the geomagnetic daily variation, but grouped the data into annual data bins thus losing information on the seasonal dependence. Thus, in theory, it is possible to extend HA-based models to take account of additional geophysical parameters; however, this would significantly add to the complexity of the process of analysis, synthesis, and interpolation. In contrast, the addition of more geophysical parameters to an ANN-based model has negligible effect on the complexity of the computational process; the only cost is in increased computing time to train the network. Besides season and latitude, geophysical parameters which are considered in our new ANN-based regional model are sunspot number and degree of magnetic activity. If deemed desirable or

necessary, the addition of further parameters, e.g. characterising ionospheric electron concentration or atmospheric tides, to the ANN-based model would be quite simple, provided the data are available. Furthermore, the extension of the geomagnetic daily variation model to cover a full hemisphere or the globe would, in computational terms, be relatively straightforward.

## 5 Development of a geomagnetic daily variation model

We designate the model which we develop as a geomagnetic daily variation model rather than as an Sq variation model. The reason is that the Sq variation is often regarded as an idealised variation which would be observed when geomagnetic disturbance effects are at an absolute minimum and computed using the mean value around local midnight as the base line level (Matsushita, 1967). On the other hand, since our geomagnetic daily variation model is intended to be used in practical applications under natural conditions, it allows for a certain degree of geomagnetic activity and uses the daily mean value as the base line level, as shown in Eq. (1).

The ANN used for the development of our geomagnetic daily variation model was a feed forward, fully connected network with one hidden layer and was trained using the back-propagation algorithm (Haykin, 1994), which continues to be the most widely used supervised training method. During this iterative process the ANN was trained to map a set of input parameters to a single target value. The input parameters which could be used were LT hour, day number of the year (which represents a smooth and continuous change in season), geomagnetic latitude of the observatory, sunspot number, and  $ak$  index representing the degree of geomagnetic activity. Both LT hour and day number suffer from the problem that the parameters for temporally adjacent data, such as for hours 24 and 1 of consecutive days, or for day numbers 365 and 1 of consecutive years, are numerically far apart. In order to enable the ANN to view such data values as being adjacent, the LT hour and day number were each split into two inputs using sine and cosine functions as described by Williscroft and Poole (1996). The target output values were the observed hourly mean values of a selected component of the geomagnetic field corresponding to the input parameters. The architecture of the neural network used for this purpose is illustrated in Fig. 2. The central goal in ANN training is not to memorize the training data, but rather to model the underlying generator of the data (Bishop, 1996). An important factor in achieving this goal is that the amount of training be sufficient, but that the network not be over trained. A recommended method to ensure that an ANN is not over trained is to stop training when the error measured using an independent validation data set starts to increase. Consequently, an independent data set was used to determine when the ANN was sufficiently trained; training ceased when the RMS error between the computed and target output values reached



**Fig. 2.** The architecture of the artificial neural network (ANN) used to develop the geomagnetic daily variation model

a stable minimum. Initially, we found that the training of the ANN took a large number of epochs and consequently also a large amount of computing time. In order to reduce these we introduced the *delta-bar-delta* algorithm (Bishop, 1996) for updating the weights. This algorithm utilises a separate learning rate for each weight in the network, with procedures for updating these learning rates during the training process.

We commenced the development and testing of the ANN-based method for generating a geomagnetic daily variation model by using very small training data sets, which were selected so that most of the geophysical parameters effectively remained constant. The training data set size, with the associated dependence on geophysical parameters, was then gradually increased as we gained a feel for the network architecture best suited to reproduce the daily variation. The number and selection of nodes in the input layer of the ANN were determined by the data set of geomagnetic field hourly mean values used for training the ANN. For example, if using a set of hourly mean values from one observatory and spanning one month only, it was assumed that the daily variation is for unchanging season, latitude, sunspot number, and level of geomagnetic activity, thus requiring only two input nodes to represent local time. At the other extreme, however, was a data set of observed hourly mean values for three observatories and spanning 12 years; in this case all input parameters were used, thus requiring seven input nodes. In all cases, observed hourly mean values were excluded from the data set if the K index was 5 or greater. The number of nodes in the hidden layer was determined by making use

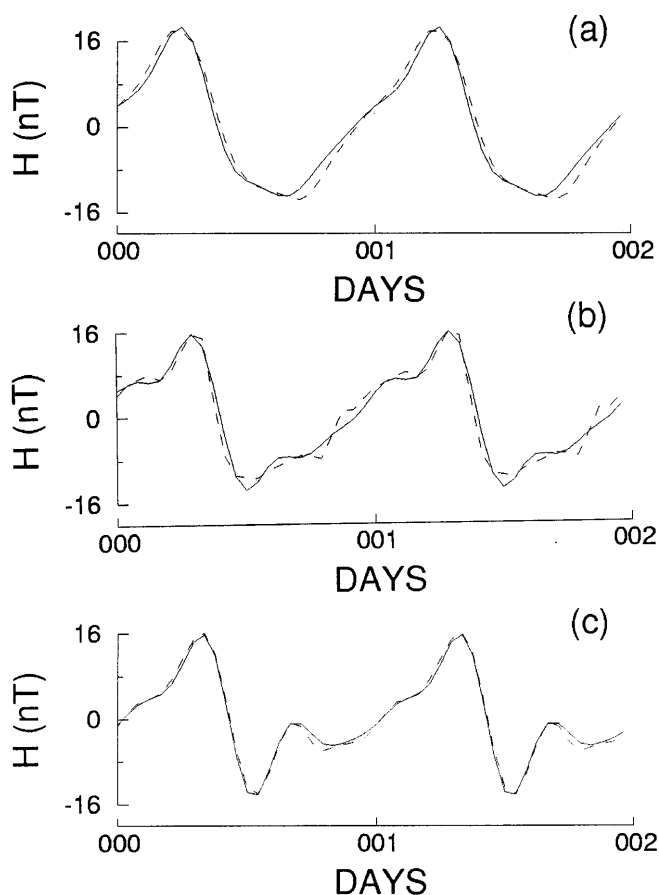
of the minimum RMS error between the computed and target output values for the independent data set. Commencing with an insufficient number of hidden nodes, we found that as the number of hidden nodes was increased, this minimum RMS error generally tended to first decrease and then to increase. The number of hidden nodes leading to the minimum value was regarded as the optimum number.

In order to assess the ability of the ANN to model the geomagnetic daily variation, the trained ANN was used to estimate the daily variations for a number of days with specified parameters. The results were then compared, either graphically or statistically, with either the observed or the harmonically synthesised daily variations for the same set of parameters. The latter daily variations were obtained by synthesis of the harmonic coefficients determined by harmonic analysis (Malin and Chapman, 1970) using a one month data set of observed geomagnetic field hourly mean values; days on which any of the K indices were 5 or greater were excluded from the analysis. For graphical assessments, the daily variations for two day intervals in the middle of a particular month, that is 14th and 15th, were generally plotted. For statistical assessments, the RMS differences between the hourly mean values of the respective daily variations were computed using data for a specified interval of time.

## 6 Results

The ANN's ability to model the geomagnetic daily variation for the simple case with constant season, latitude, sunspot number, and level of geomagnetic activity was first assessed; consequently, only two input nodes, namely those for the local time sine and cosine functions, were used. Three separate ANNs were trained using the H-component hourly mean values from Hermanus for the one month periods of January, April, and July 1993; the optimum numbers of hidden nodes were found to be 6, 20, and 10 respectively. The ANNs were then each used to estimate the daily variations for two-day periods in the middle of the same months. A comparison of the results with the daily variations determined by harmonic analysis for the same one month periods are shown in Fig. 3; the respective curves exhibit strong similarities. The RMS errors between the observed and HA daily variations and the observed and ANN daily variations respectively for each of the three one-month periods are presented in Table 2. It is seen that the differences between the two methods for computing the geomagnetic daily variation are very small. Since the HA method of computing the daily variation is well tested and accepted, these comparisons indicate that the ANN method also provides acceptable estimates of the daily variation.

We next tested the ANNs ability to model the daily variation with a seasonal dependence, but still holding the latitude, sunspot number, and level of geomagnetic activity fixed; this required two additional input nodes to represent the day number sine and cosine functions.

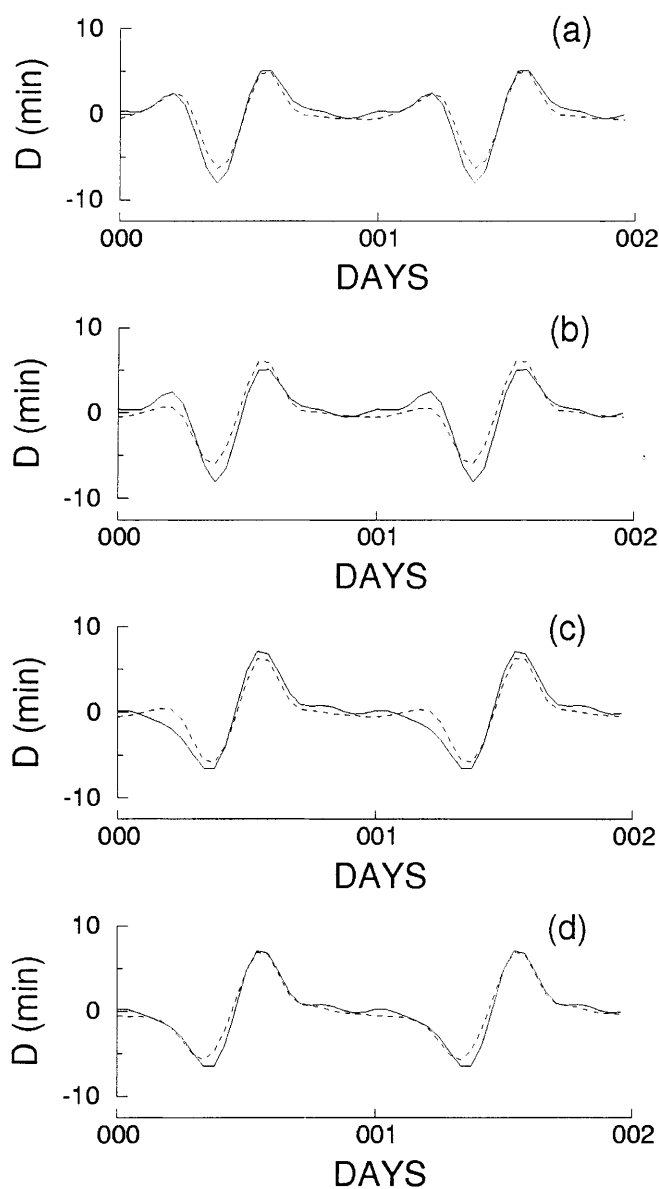


**Fig. 3a–c.** A comparison of the daily variations determined by HA (solid curves) and ANN (dashed curves) methods. Three ANNs were trained using H-component data from HER for separate one month periods during 1993. The ANNs were then each used to estimate the daily variations for two-day intervals in the middle of the same months, namely **a** January, **b** April, and **c** July

**Table 2.** RMS errors between observed H-component daily variation curves at HER and those determined by the HA and ANN methods respectively for three separate one month periods in 1993

	January	April	July
Observed-HA	9.58	8.00	6.89
Observed-ANN	9.68	8.11	6.93

The ANN was trained using the D-component hourly mean values for two years, namely 1993 and 1994, from HER; in this case the optimum network required 15 hidden nodes. The ANN and HA estimated daily variations for four two-day periods during the interval from mid September to mid October are shown in Fig. 4; the curves for the 15th of each month are very similar. This figure illustrates an advantage of the ANN determined geomagnetic daily variation curves, namely that they change smoothly from day to day, due to the seasonal dependence being determined by day number. The daily variations yielded by the HA method are the same for all days in a particular month and exhibit a discrete jump between months, due to the grouping in

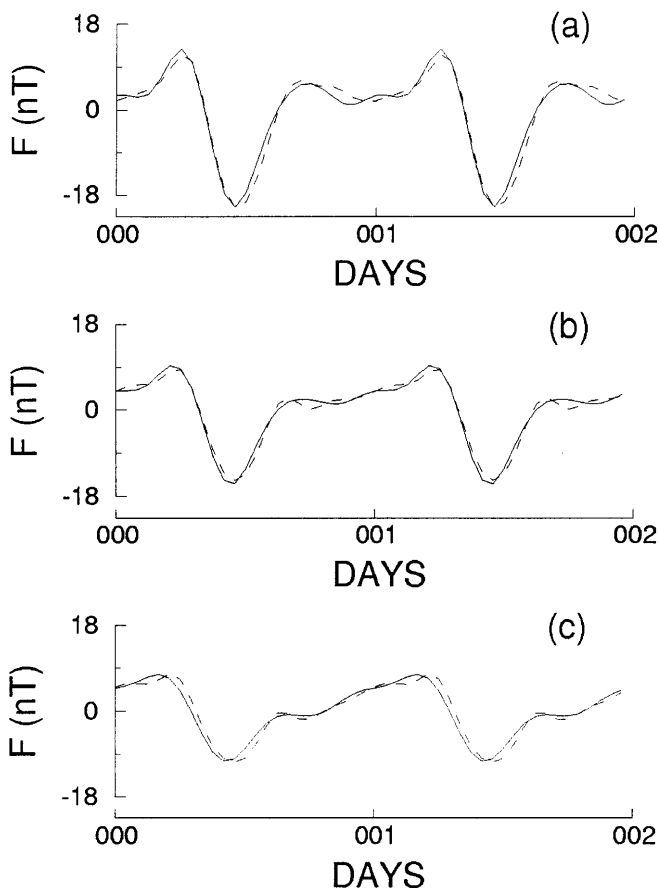


**Fig. 4a–d.** A comparison of the daily variations determined by HA (solid curves) and ANN (dashed curves) methods. The ANN was trained using the D-component data from HER for 1993 and 1994. The estimated daily variations for two-day intervals in 1993 are shown, representing the daily variations for **a** 15 September **b** 30 September **c** 1 October and **d** 15 October

monthly bins. This is clearly seen by comparing the respective curves for September 30th and October 1st. It would be necessary to introduce some sort of interpolation between monthly HA coefficients in order to obtain smoothly changing curves; Campbell and Schiffmacher (1985) achieved this by determining the mean, annual, and semiannual Fourier coefficients for each 12-member (monthly) set of HA coefficients.

Our next investigation was of the ANNs ability to model the daily variation with the addition of a latitudinal dependence, but still independent of sunspot number and level of geomagnetic activity, thus requiring five input nodes. In this case, the ANN was trained using the *F*-component hourly mean values from the

three stations HER, HBK, and TSU, for the years 1994 and 1995. It was found that an architecture with 25 hidden nodes produced a network with the minimum error. The trained ANN was then used to estimate the daily variations at two of these stations, namely HER and HBK, for a two-day interval in the middle of March 1995. In addition, the daily variation at a location halfway between them, with geomagnetic latitude  $-39.48^\circ$ , was estimated in order to gauge the ANN's performance at a latitude not included in the training data set. The ANN and HA estimated daily variations are compared in Fig. 5. The curves in Fig. 5a and c are those for HER and HBK respectively, while the curves in Fig. 5b are those for the intermediate latitude. The HA daily variation at this intermediate latitude was obtained in a manner similar to that used by Campbell and Schiffmacher (1985). They stated that it is a reasonable expectation that in quiet times the average surface field signatures of ionospheric source currents should change smoothly with latitude. Also, the effects of conducting regions in the Earth's mantle, where the induced Sq currents flow, can be presumed to change relatively slowly with latitude. Consequently, they assumed that there should be a smooth change of Sq



**Fig. 5a–c.** A comparison of the daily variations determined by HA (solid curves) and ANN (dashed curves) methods. The ANN was trained using the  $F$ -component data from three stations, namely HER, HBK, and TSU, for the years 1994 and 1995. The ANN was then used to estimate the daily variations for two-day intervals at **a** HER, **c** HBK, and **b** an intermediate latitude during mid-March 1995

Fourier coefficients with latitude and that a linear interpolation of coefficients could be made in latitude between adjacent stations. In order to do this, the expressions in Eq (1) for the  $n$ th harmonics at HER and HBK were each written in the form:

$$F_n \sin(nt + f_n) = A_n \sin nt + B_n \cos nt \quad (2)$$

The mean values of each of the coefficients  $A_n$  and  $B_n$  between HER and HBK were computed and then used to synthesise the daily variation curve for the intermediate latitude. It is seen that there is excellent agreement between the ANN and HA daily variation curves at all latitudes in Fig. 5. The stations TSU and HER are located close to the most northerly and southerly latitudes respectively of the southern African region. Consequently, we conclude that data from the three stations used should be sufficient to train an ANN for a latitude dependent geomagnetic daily variation model over this region.

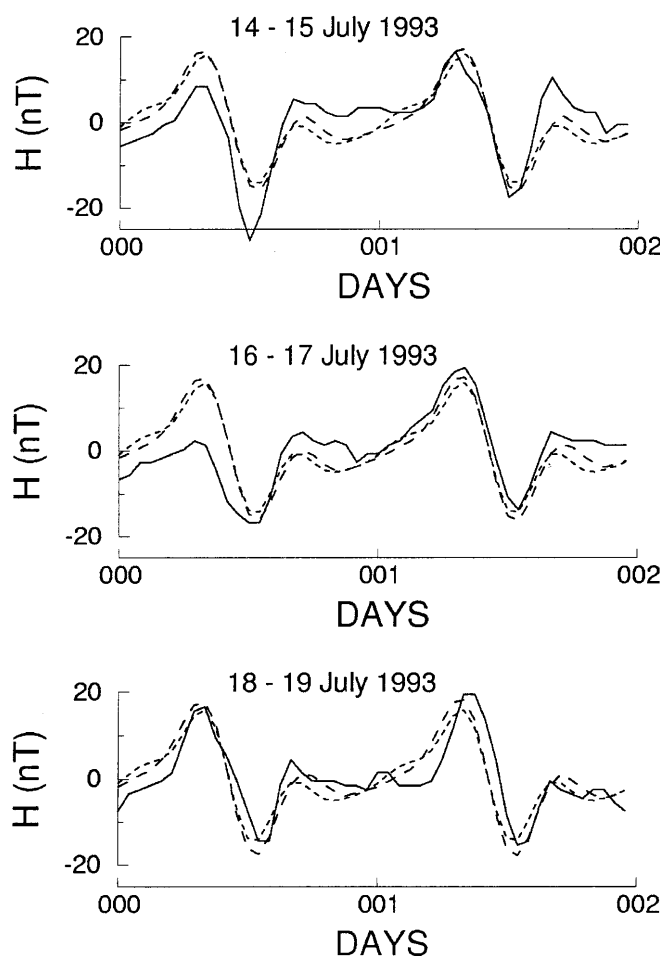
Finally, we included the sunspot number and geomagnetic activity dependencies, thus requiring seven input nodes to the ANN. Since the geomagnetic daily variation is strongly solar cycle dependent, it is essential to include this dependence in order to provide a realistic model. The dependence on geomagnetic activity is weaker, but was included primarily in an attempt to partially take account of day-to-day variability differences in the daily variation. In order to do this, the ANN was trained using the H-component hourly mean values for 12 years, namely 1978 to 1989, from the three stations HER, HBK, and TSU; the optimum number of hidden nodes was 100 in this case. When sunspot number and degree of geomagnetic activity are used as inputs to the ANN, there is a question about the optimum interval over which to average these parameters. For sunspot number we did tests using the daily values as well as averages of these values over 3, 10, 15, and 30 days. In most cases the daily values gave the minimum RMS errors, but these errors differed negligibly from those obtained using 3- or 10-day averages, indicating that the geomagnetic daily variation is not critically dependent on small changes in sunspot number. With respect to degree of geomagnetic activity, tests were done excluding the activity index (i.e. six input nodes) and also averaging the  $ak$  index over 1, 2, 4, 8, 16, and 32 three-hour intervals. In the majority of cases the values averaged over two 3 h intervals gave the minimum RMS errors; however, the differences in RMS errors obtained using these various averages were small. The reduction in RMS error achieved by the inclusion of  $ak$  averaged over two 3 h intervals compared to excluding the degree of geomagnetic activity as an input parameter was only about 5%. After training, the ANN was used to estimate the daily variations for days in various years and with various degrees of geomagnetic activity. The results were again compared with the daily variations determined by harmonic analysis using a month's data for the appropriate month and observatory.

A comparison of the observed, HA, and ANN daily variation curves for a sequence of six very quiet days, where all K indices were either 0 or 1, is shown in Fig. 6.

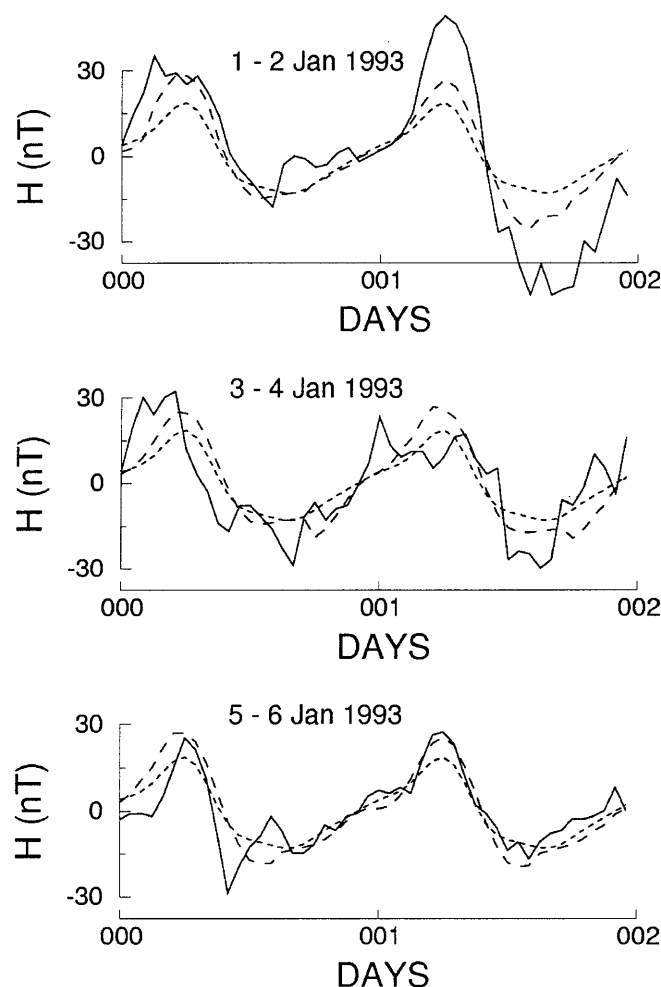
The data are for the H-component at HER for the interval 14–19 July 1993. Note that this interval does not fall within the 12 year interval used to train the ANN. Significant day-to-day variations are clearly discernable in the observed data, particularly between the 16th and 17th July, 1993, that is, the two days in the centre panel of Fig. 6. The HA curves, which were computed using the days in July 1993 with  $K$ -index  $\leq 4$ , are all identical. There is extremely little day-to-day variation between the ANN curves as a result of the small variations in sunspot numbers and  $ak$  indices. The HA and ANN daily variation curves are seen to be very similar; this similarity was confirmed by computing the RMS errors between the observed curves and the HA and ANN curves respectively, which in this case were found to be the same. Consequently, on the basis of this and other similar comparisons, we conclude that the inclusion of the degree of geomagnetic activity as an input parameter to the ANN does not contribute to the attempts to model the day-to-day variability in the case of very quiet days. Greener and Schlapp (1979) made a detailed study

of the day-to-day variability of Sq over Europe. They found that the Sq day-to-day variability arises mainly from variations in dynamo winds and electric fields. The conclusion of their study was that these are most likely driven by the diurnal evanescent (1,-1) atmospheric tidal mode.

The daily variation curves shown in Fig. 7 are similar to those in Fig. 6, except that in this case they are for a sequence of six days where the  $K$  indices were all in the range 2–4, namely, for the interval 1–6 January, 1993. The HA curves are again all identical, since they were computed using the days in January 1993 with  $K$ -index  $\leq 4$ . However, small day-to-day variations are seen in the ANN curves as the neural network attempts to compensate for the larger degree of geomagnetic activity. The RMS error between the observed and ANN values for this six day interval was 9% less than the RMS error between the observed and HA values. Thus, by introducing a dependence on geomagnetic activity, the ANN appears slightly better able to track the day-to-day variability of the geomagnetic daily variation in the case of slightly disturbed days.



**Fig. 6.** A comparison of the observed (solid curves), HA determined (short dashed curves), and ANN determined (long dashed curves) daily variations. The ANN was trained using the H-component data from three stations, namely HER, HBK, and TSU, for the 12 year period 1978 to 1989. The curves shown are for a sequence of six very quiet days, where all  $K$  indices were either 0 or 1, for the H-component at HER for the interval 14–19 July 1993



**Fig. 7.** A comparison of curves similar to those in Fig. 7, but for a sequence of six days where all  $K$  indices were in the range 2–4. The curves shown are for the H-component at HER for the interval 1–6 January, 1993



In order to obtain a clearer picture of how the HA and ANN methods for computing the geomagnetic daily variation compare, an extensive statistical comparison was made between the daily variations computed by these two methods. The ANN trained using the final data set of 12 years worth of data from three stations was used to generate the ANN geomagnetic daily variations for this comparison. Geomagnetic daily variation curves were generated for all days with  $K\text{-index} \leq 4$  in the years 1986, 1989, and 1993. The years 1986 and 1989 were sunspot minimum and maximum years respectively; they also formed part of the data set used to train the ANN. The year 1993 lay between sunspot maximum and minimum and was not part of the training data set; that is, the ANN had never seen any of the 1993 data. The HA daily variation curves used for the statistical comparisons were generated using 36 separate sets of harmonic coefficients determined from the hourly mean values grouped into separate monthly bins for each calendar month of each of the three years. The first two rows of Table 3 show the observed-HA and observed-ANN RMS errors respectively for each of the three years. The RMS errors for the ANN daily variations are very slightly larger than those for the HA daily variations, namely by 0.6%, 3.5%, and 2.4% for the years 1986, 1989, and 1993 respectively. We next computed the differences between observed-HA daily ranges and observed-ANN daily ranges. These were then grouped according to whether the HA residuals were greater than, equal to (within  $\pm 2.5\%$ ), or less than the ANN residuals. The number of days in each of these groupings are shown in the last three rows of Table 3. The results in Table 3 indicate that the geomagnetic daily variations computed using the well test HA technique are only marginally better than those obtained using the new ANN method; consequently, we conclude that the ANN method provides acceptable estimates of the daily variation under a variety of geophysical conditions.

## 7 Conclusion

There are a variety of regional and global geomagnetic field models which give the field values as functions of position and epoch. The only temporal variation generally catered for in these models is the long term secular variation; they do not provide information on short term variations such as the geomagnetic daily

variation or disturbances. There is an increasing need amongst users of geomagnetic field models for models with greater accuracy and, in particular, to have the ability to correct for short term temporal variations. Geomagnetic disturbances are likely to be one of the largest sources of deviation between observed and modelled geomagnetic field values. In recent years methods of predicting geomagnetic storms and substorms in near real-time from solar wind data using non-linear filters (Vassiliadis *et al.*, 1995) and neural networks (Lundstedt and Wintoft, 1994; Wu and Lundstedt, 1996; Weigel *et al.*, 1999) have been developed. However, such predictions are limited to hours, or at best days, into the future and require observed solar wind data. Unfortunately, accurate long term predicting or modelling of geomagnetic disturbances is not feasible at present, consequently it is difficult to correct for deviations as a result of their occurrence. Fortunately, however, large geomagnetic disturbances occur relatively infrequently. The geomagnetic daily variation, although generally smaller in magnitude than most geomagnetic disturbances, occurs on a regular daily basis. Consequently, it will be a more frequent source of deviation in any system utilising a geomagnetic field model. Furthermore, the ANN can be used to make estimates of the daily variation for any date in the past or future, provided the required input parameters (e.g. day number of the year, estimated sunspot number) are specified; consequently, these deviations can be reduced. On average one should be able to eliminate at least 50% of the geomagnetic daily variation by use of an appropriate model. Thus, for example, a geomagnetic daily variation model will find application in magnetic navigation and geophysical exploration where high degrees of accuracy are required. A geomagnetic daily variation model might also find application in reducing the daily variation effects from satellite data for main field and geomagnetic anomaly modelling.

In the past HA-based methods have traditionally been used to determine and model the geomagnetic daily variation. Here ANN-based techniques have been used for this purpose for the first time. The results presented demonstrate that an ANN can be successfully trained to model the geomagnetic daily variation regionally. The ANN method has computational advantages over the HA method. The process of computation is simplified, particularly if one wishes to take account of more variables. It is thus more amenable to computation of a realistic model, since it is easier to take account of variables such as latitude, sunspot number, and degree of geomagnetic activity. Inclusion of the latter parameter has slightly enhanced the ability to take account of the day-to-day variability in the daily variation model; however, the ability to adequately represent the day-to-day variability still remains a problem to be resolved. Applying the ANN method to other regions or even globally should not be a problem, provided the data for training the neural network are available.

*Acknowledgements.* Topical Editor K.-H. Glassmeier thanks two referees for their help in evaluating this paper.

**Table 3.** Comparison statistics between H-component daily variation curves determined by the HA and ANN methods at HER for the years 1986, 1989, and 1993

	1986	1989	1993
Observed-HA RMS errors	7.15	9.79	7.84
Observed-ANN RMS errors	7.19	10.13	8.03
Number of days on which HA > ANN	124	100	112
Number of days on which HA = ANN	36	21	22
Number of days on which HA < ANN	166	159	170

## References

- Barton, C. E., International geomagnetic reference field: the seventh generation, *J. Geomagn. Geoelectr.*, **49**, 123–148, 1997.
- Bishop, C. M., *Neural networks for pattern recognition*, Oxford University Press, New York, 1996.
- Campbell, W. H., The regular geomagnetic field variations during quiet solar conditions in *Geomagnetism*, vol 3, Ed. J. A. Jacobs, Academic Press, London, 1989.
- Campbell, W. H., and E. R. Schiffmacher, Quiet ionospheric currents of the northern hemisphere derived from geomagnetic field records, *J. Geophys. Res.*, **90**, 6475–6486, 1985.
- Campbell, W. H., and E. R. Schiffmacher, Quiet ionospheric currents of the southern hemisphere derived from geomagnetic field records, *J. Geophys. Res.*, **93**, 933–944, 1988.
- Gorney, D. J., H. C. Koons, and R. L. Waltersheid, Some prospects for artificial intelligence techniques in solar-terrestrial predictions, *Solar-Terrestrial Predictions: Workshop IV*, Eds. Hruska et al., Boulder USA, **2**, 550–564, Sept. 1993.
- Greener, J. G. and D. M. Schlapp, A study of day-to-day variability of Sq over Europe, *J. Atmos. Terr. Phys.*, **41**, 217–223, 1979.
- Haykin, S., *Neural networks. A comprehensive foundation*, Macmillan College Publishing Company, New York, 1994.
- Hibberd, F. H., Day-to-day variability of the Sq geomagnetic field variation, *Aust. J. Phys.*, **34**, 81–90, 1981.
- Joselyn, J., H. Lundstedt, and J. Trolinger, Eds., *Proce. International Workshop on Artificial Intelligence Applications in Solar-Terrestrial Physics*, Boulder USA, Dec 1993.
- Kerridge, D. J., Applications of geomagnetism in the oil industry, *Paper 05.04.02 presented at the 7<sup>th</sup> Scientific Assembly IAGA*, Buenos Aires, 1993.
- Koons, H. C., and D. J. Gorney, A neural network model of relativistic electron flux at geosynchronous orbit, *J. Geophys. Res.*, **96**, 5549–5556, 1991.
- Kotzé, P. B., Regional magnetic field modelling, *Proce. Workshop on the Hermanus Magnetic Observatory's Geophysical Services*, 49–58, 1992.
- Langel, R. A., T. J. Sabaka, R. T. Baldwin, and J. A. Conrad, The near-Earth magnetic field from magnetospheric and quiet-day ionospheric sources and how it is modeled, *Phys. Earth Planet. Int.*, **98**, 235–267, 1996.
- Lundstedt, H., Neural networks and predictions of solar-terrestrial effects, *Planet. Space Sci.*, **40**, 457–464, 1992.
- Lundstedt, H., and P. Wintoft, Prediction of geomagnetic storms from solar wind data with the use of a neural network, *Ann. Geophysicae*, **12**, 19–24, 1994.
- Malin, S. R. C., and S. Chapman, The determination of lunar daily geophysical variations by the Chapman-Miller method, *Geophys. J. R. Astron. Soc.*, **19**, 15–35, 1970.
- Malin, S. R. C., A. Cecere, and A. Palumbo, The sunspot cycle influence on lunar and solar daily variations, *Geophys. J. R. Astron. Soc.*, **41**, 115–126, 1975.
- Matsushita, S., Solar quiet and lunar daily variation fields, in *Physics of geomagnetic phenomena*, Eds. S. Matsushita and W. H. Campbell, Academic Press, New York, 1967.
- McPherron, R. L., Applications of artificial intelligence to space weather forecasting, *Solar-Terrestrial Predictions: Workshop IV*, Eds. Hruska et al., Boulder USA, **2**, 611–626, Sept 1993.
- Newitt, L. R., Practical needs of users of magnetic declination information, *Paper 05.09.07 presented at the 7<sup>th</sup> Scientific Assembly IAGA*, Buenos Aires, 1993.
- Olsen, N., The solar cycle variability of lunar and solar daily geomagnetic variations, *Ann. Geophysicae*, **11**, 254–262, 1993.
- Parkinson, W. D., *Introduction to geomagnetism*, Scottish Academic Press, Edinburgh, 1983.
- Purucker, M. E., T. J. Sabaka, R. A. Langel, and N. Olsen, The missing dimension in Magsat and POGO anomaly studies, *Geophys. Res. Lett.*, **24**, 2909–2912, 1997.
- Sutcliffe, P. R., The geomagnetic Sq variation at Marion Island, *S. Afr. J. Sci.*, **73**, 173–178, 1977.
- Sutcliffe, P. R., Substorm onset identification using neural networks and Pi2 pulsations, *Ann. Geophysicae*, **15**, 1257–1264, 1997.
- Vassiliadis, D., A. J. Klimas, D. N. Baker, and D. A. Roberts, A description of the solar wind-magnetosphere coupling based on nonlinear filters, *J. Geophys. Res.*, **100**, 3495–3512, 1995.
- Weigel, R. S., W. Horton, T. Tajima, and T. Detman, Forecasting auroral electrojet activity from solar wind input with neural networks, *Geophys. Res. Lett.*, **26**, 1353–1356, 1999.
- Williscroft, L. A., and A. W. V. Poole, Neural networks, foF2, sunspot number and magnetic activity, *Geophys. Res. Lett.*, **23**, 3659–3662, 1996.
- Wu, J.-G., and H. Lundstedt, Prediction of geomagnetic storms from solar wind data using Elman recurrent neural networks, *Geophys. Res. Lett.*, **23**, 319–322, 1996.